Rice University

(713)-527-6037

DEPARTMENT OF STATISTICS
P.O. BOX 1892
HOUSTON, TEXAS 77251-1892
scottdw@rice.edu

April 18, 1990

Dr. Julia Abrahams
Program Director, Statistics and Probability
Mathematical Sciences Division
Office of Naval Research (411-SP)
Department of the Navy
800 North Quincy Street
Arlington, VA 22217-5000

Dear Julia,

Enclosed is my final report for my ONR contract N00014-85-K-0100 for the two-year period January 1, 1988 through December 31, 1989. I believe those years have been successful and our work continues to be well-received within the statistical community.

The focus of our research has been the development of new algorithms and theory for nonparametric estimation of multivariate statistical functions, and the implementation of these algorithms in interesting computing environments. These research has involved many collaborations, including Wolfgang Härdle, Joe Austin, George Terrell, Shean Chiu, and James Thompson. I have two contract offers for a book covering this research. Two students working under my directions are completing their Ph.D. dissertations and will receive their degrees in May, 1990. Both students worked on nonparametric regression topics, one a new robust algorithm and the other a practical thesis looking at trends in ozone levels in Houston over the past decade.

The fundamental algorithm developed for multivariate density estimation has been the so-called *averaged shifted histogram*. This algorithm has been extended to other statistical functionals, for example, regression and average derivative estimates, in a joint paper with W. Härdle. This paper is expected to be read before the Royal Statistical Society in the next year.

Work on visualizing multivariate statistical functions has been presented to a wide audience at scientific meetings and at colloquiums around the country. The ideas have appeared in proceedings, a book chapter, and a book that is nearly complete. Of particular interest is a paper with Mark Hall, a graduate student in Computer Science, that appeared in the Interface of Statistics and Computer Science.

A theoretical topic that received much effort is the automatic calibration of the estimators. These cross-validation algorithms have been circulated by e-mail. Chiu has generalized the results from independent to correlated data. Scott has collaborated with Heinz-Peter Schmidt to test the algorithms on very large economic data sets from British and German household surveys. This latter work was reprinted as a book chapter by Springer Verlag.

As Associate Editor of both the top statistics journals, I have been asked to organize scientific sessions, review books, serve on NSF and NIH funding panels, and offer comments on papers. We are quite proud of the accomplishments during the relatively brief two year funding period.

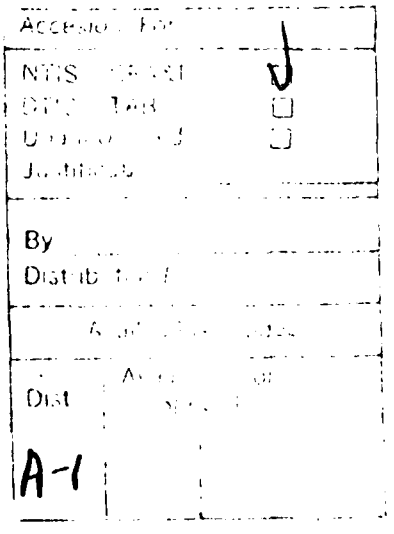
Attached is a complete list of all activities for the period covered. I would be pleased to send any selected reports you might desire. Thank you very much.

Sincerely yours,

David W. Scott

David W. Scott

cc. ONR Resident Representative, Austin



OFFICE OF NAVAL RESEARCH

PUBLICATIONS/PATENTS/PRESENTATIONS/HONORS FINAL REPORT

for

1 JANUARY 1988 THROUGH 31 DECEMBER 1989

for

CONTRACT N00014-85-K-0100

TASK NO. NR 042-535/R&T NO. 4114535

CONSTRUCTION OF A REACTIVE COMPUTING ENVIRONMENT FOR DENSITY ESTIMATION IN THREE TO FIVE DIMENSIONS

PRINCIPAL INVESTIGATOR: DAVID W. SCOTT

DEPARTMENT OF STATISTICS

RICE UNIVERSITY

HOUSTON, TEXAS 77251-1892

a. *Papers Submitted to Refereed Journals (and not yet published)*

Haerdle, W. and Scott, D.W. (1989). "Smoothing in Low and High Dimensions by Weighted Averaging Using Rounded Points," *J. Royal Stat. Society*, tentatively accepted.

Austin, J.D. and Scott, D.W. (1989). "Beyond Histograms: Averaged Shifted Histograms," Submitted to *Mathematics Teacher.*

Terrell G.R. and Scott, D.W. (1989). "A New Series for Normal Probabilities," submitted to *Probability and Statistics Letters.*

b. *Papers Published in Refereed Journals*

Chiu, S. (1988), "Weighted least squares estimators on the frequency domain for the parameters of a time series," *Annals of Statistics*, 16:1315-1326, ARO.

Chiu, S. (1989). "Detecting Periodic Components in a White Time Series," *Journal of the Royal Statistical Society, Series B*, 51:249-259.

Chiu, S. (1989). "Bandwidth Selection For Kernel Estimation With Correlated Noise." *Statistics and Probability Letters*, 8, 347-354.

Chiu, S. (1990). "Peak-Insensitive Parametric Spectral Estimation." To appear in *Stochastic Processes and Their Applications.*

Scott, D.W. (1988). "A Note on Choice of Bivariate Histogram Bin Shape." *J. of Official Statistics*, 4:47-51.

Scott, D.W. (1989). "Review of 'Nonparametric Estimation of Probability Densities and Regression Curves,' E. A. Nadaraya, author," *J. Amer. Statist. Assoc.*, to appear.

Scott, D.W. (1990). "Review of 'The New S Language,' R.A. Becker, J.M. Chambers, and A.R. Wilks, authors," *Technometrics*, 32:103-104.

Scott, D.W. and Schmidt, Heinz-Peter (1988), "Calibrating Histograms with Applications to Economic Data", *Empirical Economics*, 13:155-168.

Scott, D.W. (1988), "Comment on Paper by Haerdle, Hall, and Marron," *Journal of the American Statistical Association*, 83:96-98.

Thompson, J.R. (1989). "AIDS: The Mismanagement of an Epidemic," *Computers Math. Applic.*, 18:965-972.

c. *Books (and sections thereof) submitted for publication:*

Scott, D.W. (1990). "Some Reflections on Estimation and Visualization of Multivariate Density Functions," to appear in *Statistical Computing and Graphics Volume of the Institute for Mathematics and its Applications.*

d. *Books (and sections thereof) Published*

Scott, D.W. and Schmidt, H.-P. (1989). "Calibrating Histograms with Applications to

Economic Data," in *Semiparametric and Nonparametric Economics*, A. Ullah, Ed., Physica-Verlag, Heidelberg, pp. 33-46 (reprinted).

Scott, D.W. (1990). "Multivariate Density Estimation and Applications: Theory, Methods, and Visualization," to appear.

Thompson, J. R. (1989). *Empirical Model Building*, John Wiley and Sons.

g. *Invited Presentations at Topical or Scientific/Technical Society Conferences*

"Statistical Data Analysis," Keynote Presentation,
Thirty-Fifth ARO Design of Experiments Conference, Monterey, CA, October 18-19, 1989.

"Some Reflections on Estimation and Visualization of Multivariate Density Functions,"
Invited talk, Institute of Mathematics and Its Applications, Minneapolis, MN, August 18, 1989.

"Statistics in Motion: Where Is It Going?,"
Invited talk, ASA session "Is Real-Time Rotation Useful?", Washington, D.C., August 8, 1989.

"Visualizing Structure in High-Dimensional Data Using Density Estimation and Computer Graphics,"
Invited colloquium, Naval Postgraduate School, October 12, 1988;
Invited colloquium, Texas A&M University, October 26, 1988;
Invited colloquium, *University of North Carolina*, November 7, 1988;
Invited colloquium, University of Iowa, November 18, 1988;
Invited colloquium, University of Rochester, November 28, 1988;
Invited colloquium, Southern Methodist University, January 27, 1989;
Invited colloquium, University of Chicago, April 3, 1989.
Invited colloquium, Yale University, December 11, 1989;
Invited colloquium, Utah State University, January 25, 1990.

"Computer Graphics and Statistics,"
Invited colloquium, University Texas School of Public Health, March 20, 1989
Springwoods High School, May 11, 1989.

i. *Honors/Awards/Prizes:*

Elected President, Houston Area Chapter of the American Statistical Association, 1988-89.

Selected as Associate Editor of *The Annals of Statistics*, 1989-1991.

Selected as Associate Editor of *J. American Statistical Association*, Applications Section, 1983-85, 1986-88, 1989-91.

Member ASA Committee to Select JASA Theory and Methods Editor, 1989.

Member Advisory Committee on Continuing Education, ASA, 1988-1990.

NCI-NIH Special Review Panel, March 13-14, 1990, Bethesda, MD.

NSF SCREMS Review Panel for Special Equipment in Mathematical Sciences, March 8-10, 1989 and March 3-5, 1988.

Organizer of "Future of Statistical Computing," Statistical Computing Section Program special invited paper section for ASA150 Sesquicentennial Annual Meeting, August, 1989 (Wegman, Becker/Chambers, Thompson, Brieman).

Organizer of invited paper session entitled "Detecting Nonlinear Structure in Several Dimensions" for 1989 ASA150 meetings, August, 1989 (Flury, Inselberg, and Carr).

Chairman, Rice University Committee of Computers, 1989-90

Organizer and Co-Chairman, Conference of Texas Statisticians, COTS, Houston Hyatt, March 23-24, 1990.

j. *Technical reports published or non-refereed journals:*

Scott, D.W. (1989). "Statistics in Motion: Where Is It Going?" to appear, Proceedings of Statistical Graphics Section.

Scott, D.W. (1988). "Software for Cross-Validation of Density Estimates," Rice Technical Report 88-8-311.

Scott, D.W. and Hall, M.R. (1989). "Interactive Multivariate Density Estimation in the S Language," *Proceedings of the 20th Interface of Computer Science and Statistics*, American Statistical Association, Alexandria, Virginia, pp. 241-245.

Scott, D.W. (1987), "Software for Univariate and Bivariate Averaged Shifted Histograms," Rice Technical Report 311-87-1, ARO.

Scott, D.W. (1990) "Statistical Data Analysis: How Far Will Computer Graphics Take Us?" to appear Proceedings of the Army Design of Experiments Conference.

Thompson, J.R. (1989). "Microaxiom based simulation: a modern alternative to the closed form," *Proceedings of the American Statistical Association: Sesquicentennial Invited Paper Sessions*, Ed. M. Gail and N.L. Johnson, pp. 259-271.

k. *Graduate Students and Postdoctorals Supported Under the Contract:*

No. Graduate Students: 4
No. Postdoctorals: 0

Number of American Graduate Students: 3
Ferdie Wang, Ph.D. degree, May, 1990.
Kerry Go, Ph.D. degree expected May, 1991.
Gayle Lynn Dittrich, master's degree expected August, 1990.
Number of Foreign Graduate Students (computing support only): 1
Roland Sanchez, Ph.D. degree, May, 1990.